scientific reports

Check for updates

OPEN

Evolving evidence on a link between the ZMYM3 exceptionally long GA-STR and human cognition

H. Afshar[1], S. Khamse[1], F. Alizadeh[2], A. Delbari[1], R. Najafipour[3], A. Bozorgmehr[4], M. Khazaei[1], F. Adelirad[5], A. Alizadeh[6], A. Kowsari[7] & M. Ohadi[1]✉

The human X-linked zinc finger MYM-type protein 3 (*ZMYM3*) contains the longest GA-STR identified across protein-coding gene 5′ UTR sequences, at 32-repeats. This exceptionally long GA-STR is located at a complex string of GA-STRs with a human-specific formula across the complex as follows: (GA)8-(GA)4-(GA)6-(GA)32 (ZMYM3-207 ENST00000373998.5). *ZMYM3* was previously reported among the top three genes involved in the progression of late-onset Alzheimer's disease. Here we sequenced the *ZMYM3* GA-STR complex in 750 human male subjects, consisting of late-onset neurocognitive disorder (NCD) as a clinical entity (n = 268) and matched controls (n = 482). We detected strict monomorphism of the GA-STR complex, except of the exceptionally long STR, which was architecturally skewed in respect of allele distribution between the NCD cases and controls [F (1, 50) = 12.283; p = 0.001]. Moreover, extreme alleles of this STR at 17, 20, 42, and 43 repeats were detected in seven NCD patients and not in the control group (Mid-P exact = 0.0003). A number of these alleles overlapped with alleles previously found in schizophrenia and bipolar disorder patients. In conclusion, we propose selective advantage for the exceptional length of the *ZMYM3* GA-STR in human, and its link to a spectrum of diseases in which major cognition impairment is a predominant phenotype.

Abbreviations

AD	Alzheimer's disease
BPD	Bipolar disorder
NCD	Neurocognitive disorder
SCZ	Schizophrenia
STR	Short tandem repeat
TSS	Transcription start site
UTR	Untranslated region
ZMYM3	Zinc finger MYM-type protein 3

Human-specific characteristics and phenotypes such as late-onset neurocognitive disorder (NCD) (also known as dementia) are likely to be the consequence or by-product of human-specific evolutionary events. In agreement with the above model, recent emerging evidence indicates that signals of brain evolution in anatomically modern humans are strongly related to the Alzheimer disease (AD) pathways[1]. Remarkably, certain human-specific derived alleles protect against post-reproductive cognitive decline[2].

In comparison to single nucleotide substitutions, short tandem repeats (STRs) offer a significantly more versatile reservoir of genetic variations that may be necessary for speciation and species-specific phenotypes[3]. Following a genome-scale analysis of all human protein- coding genes annotated in the GeneCards database, we previously reported a catalog of genes containing "exceptionally long" STRs (> 5 repeats) in their core promoters[4,5]

[1]Iranian Research Center on Aging, University of Social Welfare and Rehabilitation Sciences, Tehran, Iran. [2]Department of Genomic Psychiatry and Behavioral Genomics (DGPBG), Roozbeh Hospital, School of Medicine, Tehran University of Medical Sciences (TUMS), Tehran, Iran. [3]Cellular and Molecular Research Centre, Research Institute for Prevention of Non Communicable Disease, Qazvin University of Medical Sciences, Qazvin, Iran. [4]Iran Psychiatric Hospital, Iran University of Medical Sciences, Tehran, Iran. [5]Department of Health Education and Promotion, Faculty of Health Sciences, Tabriz University of Medical Sciences, Tabriz, Iran. [6]Medical Microbiology Research Center and Microbiology Department, Qazvin University of Medical Sciences, Qazvin, Iran. [7]Health Management and Social Development Research Center, Golestan University of Medical Sciences, Gorgan, Iran. ✉email: mi.ohadi@uswr.ac.ir